

Persian Sign Language Numbers Recognition Using Thinning Method

Alaa Barkoky^{#1}, Nasrollah M. Charkari^{#2}

[#] Dept. of electrical engineering & computer science, Tarbiat Modares University, Iran

¹barkoky@modares.ac.ir; ³charkari@modares.ac.ir

Abstract-Sign language is the most natural way of communication for the people with hearing problems. One of its most appealing applications is developing a more effective interfacing of human-machine interaction. A hand gesture recognition system can provide an opportunity for deaf persons to communicate with normal people without the need of an interpreter or intermediate. In this article, we propose a method to recognize the image-based numbers of Persian sign language (PSL) using thinning method on segmented image. In this approach, after cleaning thinned image, the real endpoints have been used for recognition. The method is qualified to provide real-time recognition and is not affected by hand rotation and scaling. Experimental results on 300 images show that our approach recognition rate is 96.6 % as average.

Keywords-Persian Sign Language; Gesture Recognition; Segmentation; Static Hand Gesture Recognition; Thinning

I. INTRODUCTION

Human beings use different gestures and body movements to convey meaning and communicate with other people in their daily lives. Examples of meaningful gestures include greeting hello or goodbye using the hands, or indicating numbers using fingers.

Gesture is one of the most natural and meaningful forms in HCI¹. In other words, a gesture is a meaningful concept of motions done by human and is of supreme importance in designing an intelligent and efficient human-computer interface [1].

HCI refers to the interaction between human and computer. It is a set of interactions between human users on one hand and intelligent agent software on the other hand. This relation can be established by various members of the human body, including: hand with different state, emotional faces, movement of the human body or any combination of body organs, such as hand and arm [1].

Today, gesture recognition, and especially hand gesture recognition, has different and considerable applications, such as; vision and robotic [1-2], computer games, communication deaf people with computer [3-7], and emotional recognition based on face.

Sign language recognition is a research area involving pattern recognition, computer vision, natural language processing and psychology [3]. Today, there exist a significant number of deaf people in different countries. It is estimated to be about 30 million in the world. They usually use different movements to establish communication between each other. In general, hand movements could be to present numbers, letters, words and term indicator. In this regard, a new word should be formed from a combination of letters. In addition, it is found that many gestures depend on human cultural issues. Therefore,

it is not possible to define a common international language. For example, and consistent with the type of work in this paper, the representation of numbers 6 till 10, is different from Arabic and American sign languages than it is in PSL as shown as Fig. 1.

According to the literature, there is no certain work on the PSL numbers in HCI area. The study on hand movements of deaf people language, especially PSL, is an interesting research area. Our goal in this article is to find an approach to recognize PSL image-based numbers using thinning method and to be not affected by gesture rotation and scaling; and to make it a basis to recognize video-based ones later. Finally, the result will be a meaningful data for machine which can be treated in HCI.

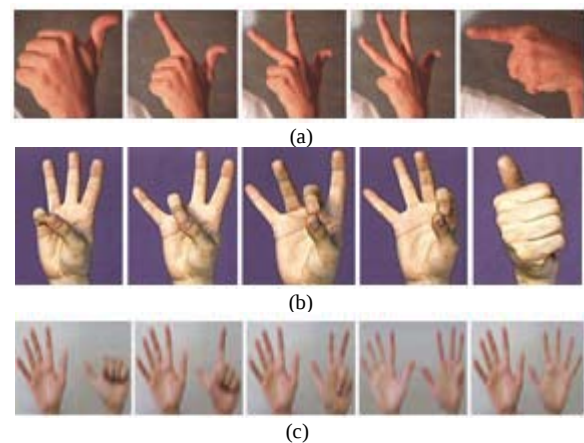


Fig. 1 Hand gestures from left to right correspond to numbers from 6 to 10, where (a) Arabic, (b) American and (c) Persian sign language.

The paper is organized as follows: section 2 briefly describes the related works; the proposed hand segmentation method is presented in section 3; section 4 explains our proposed hand gesture recognition algorithm and finally, experimental results and conclusions are presented in section 5 and 6, respectively.

II. RELATED WORKS

The Hand gesture recognition research has been grown rapidly since 2000. Hand gesture recognition has two branches: one based on static images (image-based); and the other, on dynamic sequences (video-based). Dynamic hand gestures come from frames added with time, so that they can describe complex information; so, dynamic hand gestures are not easy to recognize. Static hand recognition mainly includes the following components: segmentation, feature extraction, Recognition, as in dynamic hand gesture recognition, the object tracking, Key Frame selection and direction encoding are the significant components. In both image-based and video-based methods, some researchers used data glove of hand gesture recognition. For example, Kim et al. [19], developed a 3-D

¹. Human Computer Interaction.

hand motion tracking and gesture recognition system via a data glove. The data glove is capable of transmitting hand motion signals to a PC through wireless communication via Bluetooth. The system has been developed to recognize simple hand gestures namely scissor, rock, and paper. In general, the glove-based is a good method and has acceptable results but at the same time, it has a functional limitation.

Interesting researches have been done in this area. Control robotic is one of the most important applications. Alijanpour et al. [2], proposed a method based on hand center tracking and used correlation coefficient for state matching; for device control using inner distance feature; six gestures including of turn right, turn left, etc. has been defined on this work. Fang et al. [7] have proposed a real-time hand gesture recognition method for robot control by finding the blob and ridge of hand gesture. There are some other approaches proposed in this field, such as, Fuzzy Decision Tree [10] and Hidden Markov Model [13] to extract feature vectors and finally recognize the hand gestures.

On the other hand, some other important researchers have worked on different deaf people languages processing, Al-Jarrah and Halawani [4] presented a method for automatic translation of gestures of the manual alphabets in the Arabic sign language. They have designed a collection of ANFIS networks, each of which is trained to recognize one gesture. Incertis et al. [6] introduced an approach for deaf-people interfacing using computer vision. The recognition of static alphabetic signs of Spanish Sign Language was addressed. The proposed approach combines a number of norms to evaluate the distance of the current sign, to the sign models stored in a database. This solution leads to a largely selective criterion. Rokade et al [3] used thinning method on one to ten numbers of American Sign Language (ASL). The feature vector consists of the angles between the line, which join the center point to each endpoint, and center point to the vertical extended line. As the feature vector contains some corners, the method required vertical hand images as input. Ren and Zhang [8] presented MEB-SVM to classify gestures and finally achieve the proper recognition. Karami et al [17] proposed a system for recognizing static gestures of alphabets in PSL using Wavelet transform and neural networks. The discrete wavelet transform is applied on the gray scale images. Finally, the extracted features are used to train a Multi-Layered Perceptron neural network.

other approaches such as; local linear embedding, Neural Network shape fitting, object based key frame selection, and Haar wavelet representations have been presented in [9], [11-12] and [5].

III. HAND SEGMENTATION METHOD

The original goal of this step is to extract the hand gesture region and separating it from other objects in image. A good segmentation results give us a better accuracy rate in the next step. Pixels corresponding to the gesture are set to white and the background and other objects to black color. The all steps of hand segmentation method used in our work are plotted in Fig. 2.

A. Hand Detection

This step starts with image capturing by the camera under natural light conditions. Regardless of image size, is rescaled that considers an appropriate processing time and quality at the same time. The size which we have adopted is 150×224 pixels.

We recall that input image maybe one or two hands. One to five numbers is presented by one hand and six to ten by two hands in PSL.

Several color spaces and color-based approaches have been proposing in the literature for skin detection applications. RGB is the most commonly used color space for storing and representing digital images. A pixel (x, y) belongs to a skin if its (R, G, B) component satisfies the following conditions [16]:

$$R > 95 \text{ and } G > 40 \text{ and } B > 20 \text{ and}$$

$$\text{Max}\{R, G, B\} - \text{min}\{R, G, B\} > 15 \text{ and} \quad (1)$$

$$|R-G| > 15 \text{ and } R > G \text{ and } R > B$$

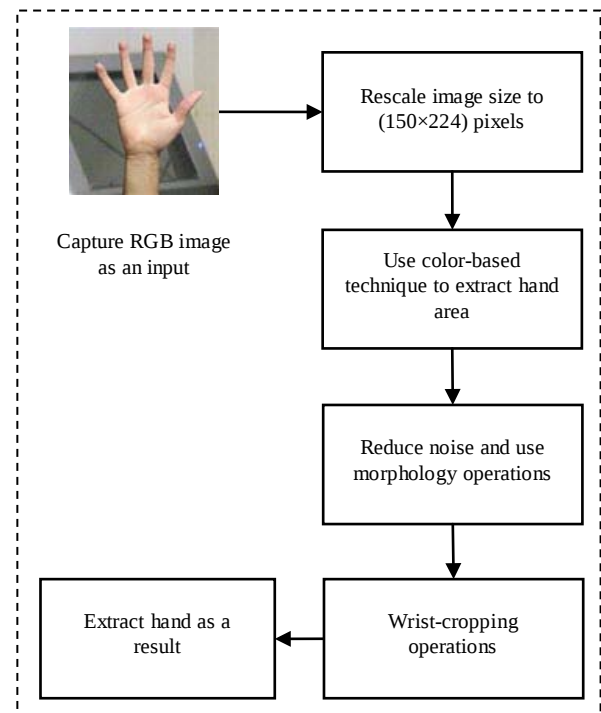


Fig. 2 Steps of hand segmentation method.

On the other hand we know that the orthogonal color spaces like YCbCr, YIQ, YUV and YIS, reduce the redundancy present in RGB color channels and represent the color with statistically independent components [15]. As the luminance and chrominance components are explicitly separated, these spaces are favorable choice for skin detection [15]. The YCbCr space represents color as luminance (Y) and chrominance (Cb and Cr) computed as a weighted sum of RGB values [14], where Y, Cb, and Cr components and values to extract human skin [14] are found by the relations “2” and “3” respectively;

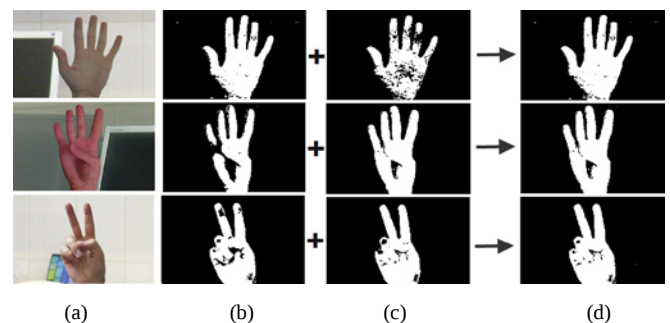


Fig. 3 (a) Original image, (b) YCbCr color space result, (c) RGB color space result and (d) result of combine both two color spaces

$$Y = 0.299R + 0.587G + 0.114B$$

$$Cr = 128 + 0.5R - 0.418G - 0.081B \quad . \square.$$

$$Cb = 128 - 0.168R - 0.331G + 0.5B$$

$$85 < Cb < 135 \text{ and } 135 < Cr < 180 \text{ and } Y > 80 \quad . \square.$$

Considering to skin color differences between humans; we have suggested to merge "use OR operation" RGB color space result with YCbCr result to extract skin area and get better segmentation results. Fig. 3 presented some results of both color spaces and combines them.

B. Reconstruction of hand area

In this step; filtering and morphology operations are performed to decrease noises and segmentation errors. A 5×5 median filter and morphology operation (close) that contains two steps, dilation and erosion is applied to do that. Fig. 4 shows the filtering and morphology results.

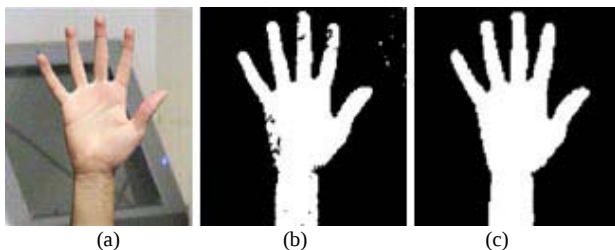


Fig. 4 (a) Original image, (b) segmented image and (b) applying filtering and morphology operations on image.

C. Wrist-cropping

Our goal at this stage is to separate the hand from arm, by a wrist cropping. Wrist cropping is essential, regardless of user shirt type, long or sleeves one, the segmented hand image includes the arm or not. The algorithm that we utilize is proposed in [18], and is based on sudden increase in the hand width from lower arm to the hand, that represents the cropped line. Fig. 5 shows the two steps of this method and the final results.

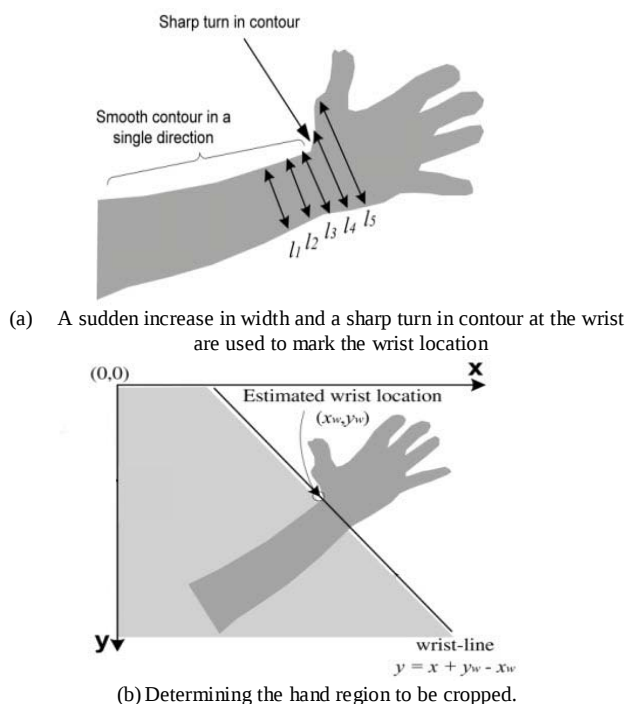


Fig. 5 A Wrist-cropping [18].

IV. HAND GESTURE RECOGNITION ALGORITHM

In this stage by getting segmented image as an input, we use the following algorithm that is shown in Fig. 6 to recognize the hand gesture.

A. Thinning and Cleaning of Segmented Image

The first step is to apply thinning method on input image. Thinning causes objects to be changed into lines. It removes pixels so that an object shrinks to a minimally connected stroke. Then, the end and joint points are found. According to 8-connectivity neighbors' pixels, the endpoint is a point that contains only one connection and represents the terminal pixel of the thin segment. Joint point is a point on the thin segment that contains more than two 8-connectivity neighbors. Joint point is the meeting point of two or more thin segments. Fig. 7 shows the mentioned points.

It is important to note that, sometimes we observe additional thin segments (non-finger thin) resulting from topography of hand or sometimes from segmentation errors. To cope with this problem, we use the following steps:

- Step 1: The length of thin segments is calculated.
- Step 2: Thin segments are sorted by length. The longest one is presented with [TMax].
- Step 3: The thin segments lengths are compared with [TMax] and those less than 40% are removed.

40% is defined as the ratio between the thumbnail and pinkie which is normally over 60%. We have used this ratio to include some special cases due to different hand sizes from one person to another. Fig. 8 shows the unclean and cleaned image.

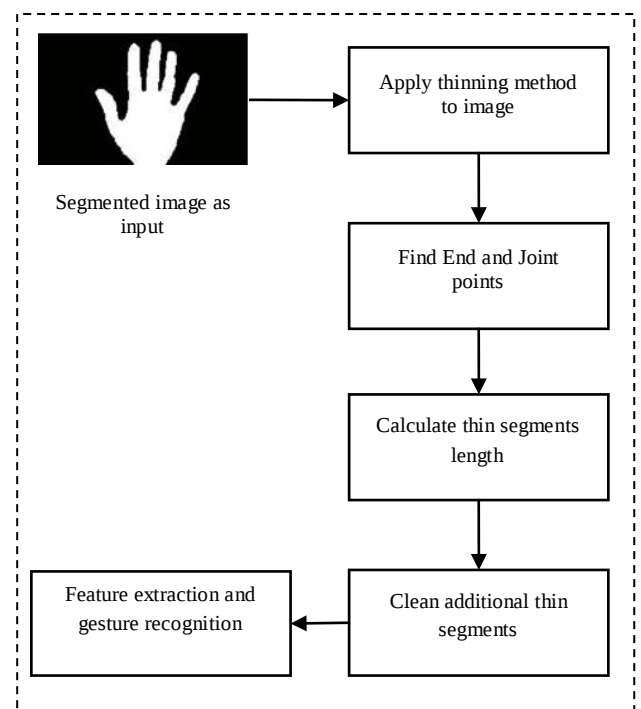


Fig. 6 Steps of hand gesture recognition algorithm.

A. Recognition of gesture

In the last step; the hand gesture will be recognized by the number of remaining endpoints in cleaned image properly. The following rule is used to recognize and distinguish ten different classes:

$$\text{If No. of endpoints in cleaned image is } \alpha \quad (4)$$

Then the gesture number is α ;

For example if the number of remaining endpoints is five then the gesture will be recognized as the number five. It is necessary to mention that, the lower section of the hand is common among almost all gestures and does not contain any useful information to distinguish itself from another. Therefore, we consider only the upper part of image to count endpoints, and the lowest joint point will be the separation point.

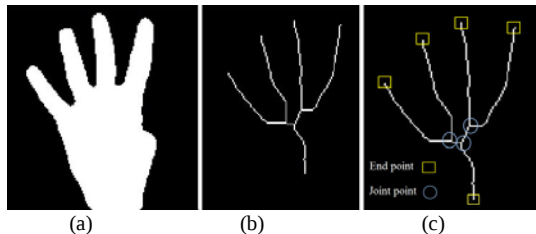


Fig. 7 (a) Segmented image, (b) Thinned image and (c) End and Joint points in thinned image.

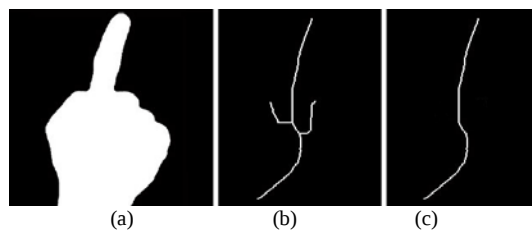


Fig. 8 (a) Segmented image, (b) thinned image with additional thin segments and (c) cleaned image.

V. EXPERIMENTAL RESULTS

The proposed system tested with 300 images, 30 samples for each number used in dataset. Considering that there is no certain dataset on PSL numbers, the dataset is made by us and we've tried to make it contain different sample models such as rescaled and rotated samples between $+45$ and -45 degree. Fig. 9 shows three different samples of number three that recognized correctly by our system. In general; the segmentation errors, recognition rate and confusion matrix for different gestures are given in Table I and Table II, respectively.

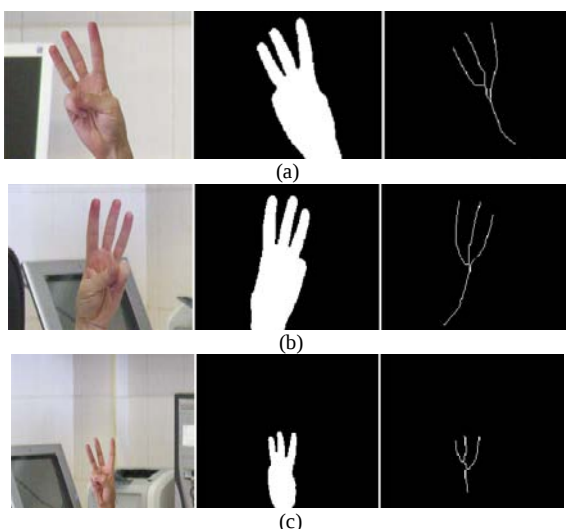


Fig. 9 Examples of (a) Rotated, (b) normal, and (c) rescaled gestures.

The average recognition rate is 96.62% which is a good result considering the diversity of data in dataset. Errors may result from the variety of factors, such as changing illumination,

background clutter, and unclear signs that lead to error in segmentation and gesture recognition later.

TABLE I EXPERIMENTAL RESULTS FOR DIFFERENT NUMBERS.

Gesture	Segmentation Error (%)	Recognition Rate (%)
One	4.9	96.6
Two	4.1	96.6
Three	2.6	93.3
Four	1.3	96.6
Five	0.6	100
Six	2.3	96.6
Seven	2.2	96.6
Eight	1.2	93.3
Nine	0.9	96.6
Ten	0.4	100
Average Recognition Rate		96.62

TABLE II CONFUSION MATRIX FOR DIFFERENT GESTURES OF NUMBERS.

Confusion matrix										
	1	2	3	4	5	6	7	8	9	10
1	29	1	0	0	0	0	0	0	0	0
2	1	29	0	0	0	0	0	0	0	0
3	0	2	28	0	0	0	0	0	0	0
4	0	0	1	29	0	0	0	0	0	0
5	0	0	0	0	30	0	0	0	0	0
6	0	0	0	0	0	29	0	0	0	0
7	0	0	0	0	0	0	29	0	0	0
8	0	0	0	0	0	0	1	28	0	0
9	0	0	0	0	0	0	0	1	29	0
10	0	0	0	0	0	0	0	0	0	30
miss	1	1	2	1	0	1	1	2	1	0

According to what we noted in the introduction that there is no certain work on the PSL numbers in HCI area, and in addition, there is difference between form gestures of PSL and other Sign languages in general, but we note some similarity between some gestures in our work and other works regardless of their meaning in sign language, therefore, we tried to compare the set of numbers (one, two, four and five) results of PSL with similar gestures in other works. The previous figure (Fig. 10) shows different classes of gestures numbers and Table. III shows comparing results with other works.

TABLE III COMPARING RESULTS WITH OTHER RESARCHES.

Research	Fingers Locating based [20]	Based on PCA approach [21]	Based on Meb-SVM [8]	Based on Thinning method [3]	Our research
Accuracy Rate (%)	91	92.36	92.57	96.65	97.45

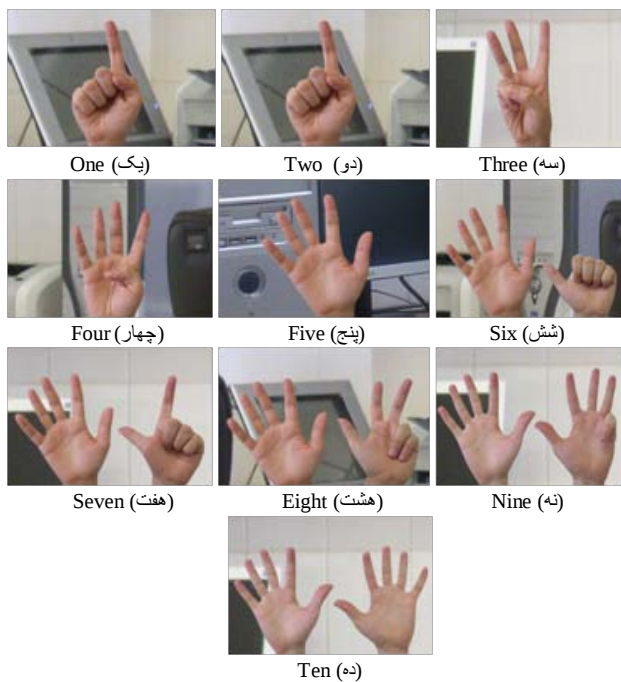


Fig. 10 Gestures of numbers 1 to 10 of Persian sign language.

VI. CONCLUSION

In this paper, we have applied a simple and fast method, which works suitably for recognize image-based numbers of PSL. In our algorithm, we utilized skin color information in segmentation phase, thinning method for feature extraction and the number of end points in cleaned image for recognition. It is a low time-consuming approach, so that a real-time recognition ratio is easily achieved. Dataset contained rescaled and rotated samples, which is considered a good advantage for our method.

This method is the base of our next work on dynamic gesture recognition which is to detect the end points Coordination in each selected key frame and track them toward hand centre of gravity of the first frame. According to this work the feature vector is created and finally given to the classifier to be recognized. In addition, we can develop the algorithm required, to include some symbol gestures such as turn right, turn left, stop, which can be used for robotic control.

ACKNOWLEDGMENT

Authors like to thanks Iran Telecommunication Research Center (ITRC) for his support of this project, and we also thank Mr. Chihada, Mr. Yousef and other image processing Lab students for helping us in preparation of hand gesture dataset.

REFERENCES

- [1] S. Mitra, Senior Member, IEEE, and T. Acharya, Senior Member, IEEE, "Gesture Recognition: A Survey," *Ieee Transactions On Systems, Man, And Cybernetics—Part C: Applications And Reviews*, Vol. 37, No. 3, May 2007
- [2] N. Alijanpour, H. Ebrahimnezhad, and A. Ebrahimi, "Inner Distance Based Hand Gesture Recognition for Devices Control " 2008 IEEE.
- [3] R. Rokade, D. Doye, and M. Kokare, "Hand Gesture Recognition by Thinning Method" *International Conference on Digital Image Processing* (2009).
- [4] Al-Jarrah, A. Halawani, "Recognition of gestures in Arabic sign language using neuro-fuzzy systems," *The Journal of Artificial Intelligence* 133 (2001) 117–138.
- [5] W. Chung, X. Wu, and Y. Xu, "A Real-time Hand Gesture Recognition based on Haar Wavelet Representation," *International Conference on Robotics and Biomimetics Bangkok, Thailand, February 21 - 26, 2009*.
- [6] Incertis, J. Bermejo, and E. Casanova, "Hand Gesture Recognition for Deaf People Interfacing," *The 18th International Conference on Pattern Recognition*, 2006 IEEE.
- [7] Y. Fang, K. Wang, J. Cheng, and H. Lu, "A Real-Time Hand Gesture Recognition Method," *Icm 2007*.
- [8] Y. Ren, F. Zhang, "Hand Gesture Recognition Based on MEB-SVM" *International Conferences on Embedded Software and Systems* (2009).
- [9] X. Teng, B. Wu, W. Yu, and C. Liu, "A hand gesture recognition system based on local linear embedding," *Journal of Visual Languages and Computing* 16 (2005) 442–454.
- [10] M. Jeon, S. Yang, and Z. Bien, Fellow, IEEE, "User Adaptive Hand Gesture Recognition using Multivariate Fuzzy Decision Tree and Fuzzy Garbage Model" *FUZZ-IEEE 2009*.
- [11] Stergiopoulou, N. Papamarkos, "Hand gesture recognition using a neural network shape fitting technique," *Engineering Applications of Artificial Intelligence Journal* (2009).
- [12] U. Rokade, D. Doye, and M. Kokare, "Hand Gesture Recognition Using Object Based Key Frame Selection," *International Conference on Digital Image Processing* (2009).
- [13] M. MONI, A. Shawkat Ali, "HMM based Hand Gesture Recognition: A Review on Techniques and Approaches," 2009 IEEE.
- [14] V. NABIYEV, A. GÜNAY, "Towards A Biometric Purpose Image Filter According To Skin Detection," *The Second International Conference "Problems of Cybernetics and Informatics"* September 10-12, 2008.
- [15] P. Kakumanu, S. Makrogiannis, and N. Bourbakis, "A survey of skin-color modeling and detection methods," *Pattern Recognition Journal* 40 (2007) 1106-1122.
- [16] V. Vezhnevets, V. Sazonov, and A. Andreeva, "A Survey on Pixel-Based Skin Color Detection Techniques," *Proc. Graphicon CONFERENCE-2003*.
- [17] Karami, B. Zanj, and A. Sarkaleh, "Persian sign language (PSL) recognition using wavelet transform and neural networks" (*Expert Syst Appl*) 38 (2011) 2661–2667.
- [18] Wah Ng, S. Ranganath, "Real-time gesture recognition system and application," *Image and Vision Computing* 20 (2002) 993–1007.
- [19] J.H. Kim, N.D. Thang, T.S. Kim, "3-D hand motion tracking and gesture recognition using a data glove," 978-1-4244-4349-9/09/\$25.00 2009 IEEE.
- [20] Malima, E. Ozgur, M. Çetin, "A fast algorithm for vision-based hand gesture recognition for robot control" 1-4244-0239-5/06 2006 IEEE.
- [21] P. Sheng-Yu, H. Kanoksak, L. Kuan-Ching, "A Real-Time Hand Gesture Recognition System for Daily Information Retrieval from Internet" *Fourth International Conference on Ubi-Media Computing*, 978-0-7695-4493-9/ 2011 IEEE.